



Federal Identity Theft Task Force
Federal Trade Commission
Office of the Secretary
Room H-135 (Annex N)
600 Pennsylvania Avenue, N.W.
Washington, D.C. 20580

As Chair of the U.S. Public Policy Committee of the Association for Computing Machinery (USACM), I welcome the request for comments by the Federal Identity Theft Task Force. Identity theft is a growing trend affecting thousands of Americans each year. It requires a national discussion at the highest levels of government in order to devise appropriate policies to reduce its occurrence. Members of our committee have been involved in issues of digital identity, information security, cyber crime investigation, and other topics for many years. Recently I provided expert testimony before the committee investigating the 2006 Veterans' Administration data breach. Our experience and study has shown us that where identity theft is concerned, two major issues go hand in hand: computer security and privacy.

Well-publicized instances of personal data exposures and misuse have demonstrated the threat of identity theft and corresponding challenges to the adequate protection of privacy. Personal data -- including copies of video, audio, and other surveillance -- needs to be collected, stored, and managed appropriately throughout every stage of its use by all involved parties. However, protecting private, personal data requires more than simply ensuring effective data security. It requires approaching personal data as a steward rather than as a custodian.

As we have detailed in our statement on privacy (<http://www.acm.org/usacm/Issues/Privacy.htm>, and included with these comments), a holistic, proactive approach to ensuring privacy is necessary, and is an important part of helping minimize the risk of identity theft. This approach gives people more control over their personal data and has the salutary effect of enhancing the early discovery of identity theft. Following the guiding principles of data minimization, consent, openness, access, accuracy, security, and accountability and our associated recommendations will go a long way toward ensuring privacy of stored data and reducing the risk of identity theft. It would help make data custodians into data stewards.

DATA SECURITY AND DATA BREACH NOTIFICATION POLICIES

A uniform national policy could harmonize company practices for protecting personal data across the United States, but such a policy, if it is not sufficient for protecting personal information, could also undermine consumer protection. We recommend that strong national standards be developed that are based on widely-accepted international data security standards. For example, ISO 17799 on information security management and ISO 18033 on data encryption are comprehensive and detailed security standards that have been adopted by the international community. These are the basis for developing data security plans, but ultimately any data security standards should be technology-neutral and based on the highest possible protections for personal data. Further, notification is often an effective method for ensuring that companies continually improve their security practices. Clearly if there is a breach, regardless of the risk to consumers, a company's security system should be hardened to deal with the vulnerabilities. A national breach notification standard could provide more transparency about ineffective or effective security practices. Last year Congress tried to establish a national breach standard based on varying degrees of risk. At that time, we expressed concern that a risk-based standard would not provide the level of transparency necessary to ensure protection of personal data. In short, implementing lowest common denominator standards for data security or notification is inefficient and raises the possibility of doing more harm than good.

USE OF SOCIAL SECURITY NUMBERS

Regarding the use of Social Security numbers (SSNs) in data records, the use of SSNs is risky and can cause (and has caused) problems for protecting privacy and reducing identity theft. If they are to be used, there should be clear guidelines limiting and auditing access to this data. Furthermore, SSNs should be used only as an identifier, and not for authentication. Possession of a SSN is not enough to confirm, or authenticate, that a person is the individual assigned to that SSN. Simply because someone has a security pass (identification) does not mean that they are the person who was issued that security pass. They would be authenticated when an authorized agent confirmed that the person presenting the pass is the person assigned to that pass.

Another way to limit access to the SSN (and by extension minimize the risk of identity theft) would be to store the SSN in a separate file linked by some unique, generated number. This decreases the relative risk in holding SSN's in databases by placing it in a less-frequently accessed file, which can also have separate, more stringent, access controls than other personal information in the database. This same practice should apply to any substitutes for a SSN, if they are used in the same ways, and with the same frequency, that SSNs are currently used.

In addition to our recommendations on privacy, we are enclosing a fact sheet on identity and authentication. The general lack of understanding of these principles has indirectly led to many of the instances of identity theft and privacy exposure. We recommend that

the Task Force be clear in its use of terms, and help to educate companies and government about these concepts.

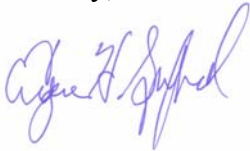
NATIONAL IDENTITY FILES

While the development of national identity files to help restore identity to victims of identity theft has a valuable policy goal in mind, it appears to pose the same kind of risks for identity theft faced by other databases. If this idea were to move forward, it should be limited to those individuals who have had their identities stolen already. Another concern is how to authenticate victims of identity theft over a long period of time.

As the attached fact sheet indicates, it can be difficult to authenticate people based on personal data. With such a system it is important to be very clear on the risks and resources involved, or these files may become another target for identity theft - a very tempting one as these files are intended to serve as the ultimate verification of identity.

Thank you for considering our views. The work of the Task Force is an important step toward increased efforts that help reduce identity theft and encourage more secure, private and reliable computer information. If USACM can provide any clarification to these comments, or answer any other technical questions, please do not hesitate to contact our Public Policy Director, Cameron Wilson, at 202-659-9711 or cameron_wilson@acm.org.

Sincerely,



Eugene Spafford, Ph.D.

Chair

U.S. Public Policy Committee of the Association for Computing Machinery

About ACM and USACM

With over 80,000 members worldwide, The Association for Computing Machinery is an educational and scientific society focused on advancing computing as a science and a profession. USACM serves as the focal point for ACM's interaction with U.S. government organizations, the computing community, and the U.S. public in all matter of U.S. public policy related to information technology.



Understanding Identity and Identification

Professionals who work with issues of security and control use some terms to precisely describe access to resources and naming. These same terms have usage in general language, but the words frequently are used imprecisely and even misleadingly. When describing how information systems operate, and when formulating regulations or laws, it is important that these terms are understood and used precisely.

The purpose of this short document is to describe these important terms for readers who are not familiar with the more formal definitions. These related terms are *identification*, *authentication*, and *authorization*. Related concepts include *uniqueness* and *biometrics*.

Terms

Identification is associating a distinguishing label (*identifier*) with something within a specific group or context. You can identify someone by getting both their label and the context of that label. An ID card can provide both the name (e.g. “John Smith”) and the context (e.g., “licensed driver”). Identification can also occur by providing only the context or group name, such as identifying oneself as a police officer, a student, a graduate of West Point, or a member of Congress by wearing an appropriate badge, uniform, or class ring. The reliability of an identification depends on the confidence that the distinguishing label and context actually apply to the individual in question.

Note that even when identification is reliable – and it often is not – it does not imply anything beyond being able to distinguish among items or people. Identification can be used to determine if someone is a member of a group or not, or among members of the group. If someone were to identify herself as “Snow White,” that is an identification if she uses it consistently. In the context of a Halloween party or an Internet chat room, that may be a logical label to adopt.

A key concept is that identification does not need to be a standard name. It can be a nickname, a login, or a simple description, such as “I am the tallest one here” or “I am the one with red hair.” Those are means to distinguish one person from another in a particular group context.

People are most often identified in social situations by their names. In the United States, these names are usually composed of a first (given) name, one or more middle names (usually), and a last (family) name. In other countries, names may be a single word, or everyone may have a common family or middle name.

Understanding Identity and Identification

Page 2

Uniqueness is when multiple items do not have the same identifier. Human names are seldom unique across a large enough population. For instance, there are many, many people named “John Smith” in the USA. If we also consider ancestors, then there may be even more individuals who have been associated with the same identifier (name).

We can further qualify an identifier to make it more specific and less likely to be a duplicate of another identifier. For instance, someone could be “John Smith who was born April 1, 1952 in Boise and whose mother was named Matilda.” However, we cannot always be certain this is unique, and it is unwieldy to use in formal documents. Thus, we commonly use an artificial identifier that is generated and assigned in a manner that ensures that it is unique within context. For instance, Social security numbers are supposed to be assigned without reuse, making them theoretically unique. Other identifiers (e.g., driver’s license numbers) are similarly generated to provide uniqueness.

Authentication is the process of verifying – to some desired level of confidence – that a claimed identifier is valid and actually associated with a particular item or person. Often, this validation is performed by one or more persons inspecting the identification and authenticator(s). The authenticators can also be examined by some technical means, such as a login program.

Authenticators of people are typically some combination of “something known,” “something possessed,” or “something about (structural)” the person. These items have been previously registered with the persons or organizations performing the authentication.

Something known is a secret or a fact that is unlikely to be known to an impostor. Passwords, when properly chosen and protected, are this form of authenticator. In many old combat movies, the spy is exposed because he doesn’t know which team won the World Series the previous year – this is another form of “something known” as a group authenticator. Many companies use items such as “mother’s maiden name,” “birth date” or “social security number” as authenticators, but this is bad practice as those items are often easily discovered facts. Many of these items are public information as a matter of law or custom.

Something possessed is a distinguishable token or a key that matches a counterpart. A license issued by a government agency is a form of token. Another example from an old movie is the dollar bill or playing card that is ripped raggedly in half – the two halves are kept and joined together to *mutually authenticate* two parties.

Something about (structure) when applied to a person is known as **biometrics**. We examine something physical about the person we wish to identify, such as a fingerprint, pattern of blood vessels inside the eye, or DNA markers. A current location may also be used for authentication, such as GPS coordinates, telephone caller-id or computer network address.

Using a combination of authenticators is known as *multi-factor authentication*.

Authorization is the granting of rights (verb) or the grant itself (noun). Generally, one authorizes an authenticated party. *Permission* is used by some people as a synonym for authorization.

Understanding Identity and Identification

Page 3

An example

Consider a scenario involving a person who wishes to enter a guarded building. When the person approaches the building to enter, a guard stops him to verify that he can enter.

The person produces an *identification* card (something possessed) issued by a trusted authority (the context). The guard compares the picture on the ID with the face of the person (a biometric) and determines that the person is the one identified by the card. She has been instructed that anyone with a valid blue card is allowed to enter, but without a cell phone, so she allows the person to pass after determining that he does not have a cell phone.

Note that this is use of multi-factor authentication, and the identification is based on group membership (“people with a valid blue badge”) – no specific name or ID number is required. Permission to enter is the authorization involved. A further element of access control that is not based on identity or authentication is also involved: there is no authorization to carry a cell phone in.

There are many potential weaknesses in this system as described. The system can be redesigned to prevent the weaknesses, but defensive measures may be too expensive or cumbersome to be worth the effort given both the likelihood of the threats occurring and the value of what is behind the door. Examples of weaknesses include:

The ID card could be faked so that a false positive authentication is made and an unauthorized person is allowed to enter.

The picture on the card may be old and the guard makes a false negative authentication: she refuses to allow the authorized person to pass.

The guard may be overpowered or bribed so that unauthorized people enter.

The guard is unable to recognize a disguised cell phone.

Someone pretending to be a law enforcement officer, in uniform, orders the guard to let him pass or he will arrest her for obstructing justice. She complies.

If too many people arrive in a short time, the guard may not be able to process them in a timely fashion.

The guard may fall ill and leave her post, leaving the door locked or unlocked for subsequent visitors.

Additional Notes

It is possible to have authentication and authorization without specific identification. For instance, producing an *authentic* \$20 bill provides authorization to make a purchase for something up to \$20 in cost. It is not a requirement to *identify* the purchaser beyond

Understanding Identity and Identification

Page 4

being a member of the group who has cash.

Knowing precise, authentic identity **does not disclose intent!** Knowing the name of everyone who enters a building or boards a plane does not mean that they will be well-behaved. Most identification checks instituted in the wake of 9/11 perform at most a weak security function because there is poor (or no) authentication, and even when the identity is known it does not prove anything about intent.

Social security numbers are not supposed to be reused. However, numerous recorded cases of SSN duplication make the use of these numbers as unique IDs problematic.

Most biometrics have been developed and tested for authentication of a claimed identity, not for performing the identification itself; fingerprints are a notable exception. Insufficient experience has been gained with many biometrics to know error rates over large populations. By example, given the biometric data that John Smith is 6'1" tall, has brown hair and green eyes, we can determine with some confidence whether a person in the room claiming to be John is actually John. However, given that same information and a crowd of people in a football stadium, we cannot be certain that we can uniquely identify John if he is present. Almost certainly, we will also make many false positive identifications.

We know that every biometric has deficiencies. Not everyone has valid fingerprints over their entire lives, height and weight change, twins and triplets have the same DNA, and so on. People with special interests in some technologies have made unsupported claims about the performance of certain biometrics.

Most organizations use weak authenticators. In part, this is because most people are poor at remembering items such as long passwords and multiple ID numbers. As noted, use of authenticators such as mother's maiden name, social security number, or other such items is poor practice because those items can be easily found for many people.

Every instance where identifiers and authenticators are to be used should be carefully analyzed to determine strengths and weaknesses. This includes the value of what is being protected, and the consequences of false positives (authenticating an incorrect identity) and false negatives (failing to authenticate a valid identity).



USACM

The Public Policy Committee of ACM

USACM Policy Recommendations on Privacy
June 2006

BACKGROUND

Current computing technologies enable the collection, exchange, analysis, and use of personal information on a scale unprecedented in the history of civilization. These technologies, which are widely used by many types of organizations, allow for massive storage, aggregation, analysis, and dissemination of data. Advanced capabilities for surveillance and data matching/mining are being applied to everything from product marketing to national security.

Despite the intended benefits of using these technologies, there are also significant concerns about their potential for negative impact on personal privacy. Well-publicized instances of personal data exposures and misuse have demonstrated some of the challenges in the adequate protection of privacy. Personal data — including copies of video, audio, and other surveillance — needs to be collected, stored, and managed appropriately throughout every stage of its use by all involved parties. Protecting privacy, however, requires more than simply ensuring effective information security.

The U.S. Public Policy Committee of the Association for Computing Machinery (USACM) advocates a proactive approach to privacy policy by both government and private sector organizations. We urge public and private policy makers to embrace the following recommendations when developing systems that make use of personal information. These recommendations should also be central to any development of any legislation, regulations, international agreements, and internal policies that govern how personal information is stored and managed. Striking a balance between individual privacy rights and valid government and commercial needs is a complex task for technologists and policy makers, but one of vital importance. For this reason, USACM has developed the following recommendations on this important issue.

RECOMMENDATIONS

MINIMIZATION

1. Collect and use only the personal information that is strictly required for the purposes stated in the privacy policy.
2. Store information for only as long as it is needed for the stated purposes.
3. If the information is collected for statistical purposes, delete the personal information after the statistics have been calculated and verified.
4. Implement systematic mechanisms to evaluate, reduce, and destroy unneeded and stale personal information on a regular basis, rather than retaining it indefinitely.
5. Before deployment of new activities and technologies that might impact personal privacy, carefully evaluate them for their necessity, effectiveness, and proportionality: the least privacy-invasive alternatives should always be sought.

CONSENT

6. Unless legally exempt, require each individual's explicit, informed consent to collect or share his or her personal information (*opt-in*); or clearly provide a readily-accessible mechanism for individuals to cause prompt cessation of the sharing of their personal information, including when appropriate, the deletion of that information (*opt-out*). (NB: The advantages and disadvantages of these two approaches will depend on the particular application and relevant regulations.)
7. Whether opt-in or opt-out, require informed consent by the individual before using personal information for any purposes not stated in the privacy policy that was in force at the time of collection of that information.

OPENNESS

8. Whenever any personal information is collected, explicitly state the precise purpose for the collection and all the ways that the information might be used, including any plans to share it with other parties.
9. Be explicit about the default usage of information: whether it will only be used by explicit request (*opt-in*), or if it will be used until a request is made to discontinue that use (*opt-out*).
10. Explicitly state how long this information will be stored and used, consistent with the "Minimization" principle.
11. Make these privacy policy statements clear, concise, and conspicuous to those responsible for deciding whether and how to provide the data.
12. Avoid arbitrary, frequent, or undisclosed modification of these policy statements.
13. Communicate these policies to individuals whose data is being collected, unless legally exempted from doing so.

ACCESS

14. Establish and support an individual's right to inspect and make corrections to her or his stored personal information, unless legally exempted from doing so.
15. Provide mechanisms to allow individuals to determine with which parties their information has been shared, and for what purposes, unless legally exempted from doing so.

16. Provide clear, accessible details about how to contact someone appropriate to obtain additional information or to resolve problems relating to stored personal information.

ACCURACY

17. Ensure that personal information is sufficiently accurate and up-to-date for the intended purposes.
18. Ensure that all corrections are propagated in a timely manner to all parties that have received or supplied the inaccurate data.

SECURITY

19. Use appropriate physical, administrative, and technical measures to maintain all personal information securely and protect it against unauthorized and inappropriate access or modification.
20. Apply security measures to all potential storage and transmission of the data, including all electronic (portable storage, laptops, backup media), and physical (printouts, microfiche) copies.

ACCOUNTABILITY

21. Promote accountability for how personal information is collected, maintained, and shared.
22. Enforce adherence to privacy policies through such methods as audit logs, internal reviews, independent audits, and sanctions for policy violations.
23. Maintain *provenance* — information regarding the sources and history of personal data — for at least as long as the data itself is stored.
24. Ensure that the parties most able to mitigate potential privacy risks and privacy violation incidents are trained, authorized, equipped, and motivated to do so.

USACM does not accept the view that individual privacy must typically be sacrificed to achieve effective implementation of systems, nor do we accept that cost reduction is always a sufficient reason to reduce privacy protections. Computing options are available today for meeting many private sector and government needs while fully embracing the recommendations described above. These include the use of de-identified data, aggregated data, limited datasets, and narrowly defined and fully audited queries and searches. New technologies are being investigated and developed that can further protect privacy. USACM can assist policy-makers in identifying experts and applicable technologies.

For more information about USACM, please contact the ACM Office of Public Policy at (202) 659-9711 or see <<http://www.acm.org/usacm/>>.